1. **Introduction:**

Kala-azar is a parasitic disease caused by "Leishmania donovani" transmitted by sand phlebotomies argentines. The disease is prevalent among socio-economically poorer sections of the society living in rural areas of Nepal. The person infected with kala-azar suffers from recurrent fever, loss of appetite, loss of weight and progressive enlargement of spleen. The disease is chronic and if not treated, it leads to death.

Visceral Leishmaniasis (VL), also known as kala-azar, black fever, and Dumdum fever is the most severe form of Leishmaniasis. Leishmaniasis is a disease caused by protozoan parasites of the Leishmania genus. This disease is the second-largest parasitic killer in the world after malaria, responsible for an estimated 500,000 infections each year worldwide. The parasite migrates to the internal organs such as liver, spleen (hence “visceral”), and bone marrow, and, if left untreated, will almost always result in the death of the host. Signs and symptoms include fever, weight loss, mucosal ulcers, fatigue, anemia, and substantial swelling of the liver and spleen. Of particular concern, according to the World Health Organization (WHO), is the emerging problem of HIV/VL co-infection.

Two species of Leishmania are known to give rise to the visceral form of the disease. The species commonly found in East Africa and the Indian subcontinent is L. donovani and the species found in Europe, North Africa, and Latin America is L. infantum, also known as L. Chagasi.

Visceral Leishmaniasis (VL), also known as Kala azar or black fever, is a vector-borne parasitic disease. Kala azar is one of the World’s most neglected and poverty-related diseases, affecting the poorest people in developing countries associated with malnutrition, weakness of the immune system, displacement, poor housing, illiteracy, gender discrimination, and lack of resources. Kala azar is spread over a large geographical area across the globe with estimated yearly incidence of Kala azar of 500,000 cases, which lead to loss of nearly 2.4 million disability-adjusted life years (DALYs) each year. Since the last century several outbreaks of Kala azar have affected South-East Asia. Various programmers to control the disease have failed despite considerable efforts being done for prevention and control of the disease.
Currently several factors have made a smooth way for achieving the goal of eliminating Kala azar from the South East Asia Region (SEAR)\textsuperscript{31}. In SEAR, the only causative agent is protozoan parasite Leishmania Donovani, the only known vector is the female Sandfly of genus Phlebotomus Argentipes and the only reservoir is human\textsuperscript{51}. Availability of effective and relatively safer drugs like Miltefosine and liposomal Amphotericin for treatment, indoor residual spray and availability of a rapid diagnostic test (rk39) makes Kala azar a candidate for elimination. India, Bangladesh and Nepal have demonstrated strong political will and commitment towards elimination of Kala azar by signing a Memorandum of Understanding (MoU) during the World Health Assembly in Geneva in 2005. The goal of Tripartite MoU was to contribute for improving the health status of vulnerable groups and at-risk population living in Kala azar endemic areas of Bangladesh, India and Nepal by the elimination of Kala azar so that it would no longer be a public health problem.

Kala azar is endemic in 88 countries, particularly in Africa, Latin America, South and Central Asia, the Mediterranean basin and the Middle East with over 200 million people at risk. In the South-East Asian Region Kala azar is prevalent in India, Bangladesh and Nepal and with few foci in Bhutan. It is endemic in 115 districts spread over this four countries namely India (52), Bangladesh (45), Nepal (12) and Bhutan (6). Epidemiologically, Kala azar follows an epidemic cycle resurging almost every 15–20 years\textsuperscript{2}. The population at risk is nearly 65 million, 6 million, and 2.5 million in Bangladesh, Nepal and Bhutan respectively\textsuperscript{34}. Regardless of lots of strategy for elimination estimated annual incidence of Kala azar cases are varying from 20 – 25 in India, 5 – 8 in Nepal and 13 –31 in Bangladesh per 10,000 population\textsuperscript{53}.

Kala azar is a major public health problem in Nepal. Altogether 12 districts in eastern end central Terai regions bordering the northern states of India are endemic. Since Kala azar in Nepal as in the rest of the World is a disease of poverty, only small proportion of the Kala azar affected population of the lowest socio economic strata living in remote rural regions has access to health institutions. Household with damp earthen floors are ideal breeding sites for sandfly, poor families are more vulnerable to the disease and less likely to seek for health care in a timely manner if sick. Therefore, the government of Nepal set a target to reduce incidence of Kala azar less than 1 case per 10,000 population at district level by end the of 2015.
The burden of Kala azar is grossly underestimated by the health systems in the Indian subcontinent due to over-reliability on passive surveillance\textsuperscript{56}. Though Kala azar patients ultimately report to health centers and hospitals, diagnosis is often missed in the early stages of infection and delayed due to lack of diagnostic facilities at peripheral levels of the health system with a consequent delay in treatment and sustaining the human reservoir. Besides this, unavailability of noninvasive rapid diagnostic test aggravates the delay the diagnosis\textsuperscript{51}. Many Kala azar cases go undiagnosed in a web of unqualified medical practitioners which further adds to considerable delay in diagnosis and treatment. Patients who are going to private sector for treatment usually stop in between due to high cost, leading to prolonged transmission\textsuperscript{69}. Strategies are to improve reporting by strengthening diagnosis and treatment besides involving the private sector. To give pace to the elimination, Kala azar could be addressed as notifiable disease\textsuperscript{42}. This will not only help to estimate the true burden of Kala azar but also hasten the treatment of patients identified with better monitoring. There is a wide gap in training and support of staff with technical and management expertise. As prevention and control activities for Kala azar are increasing in Indian subcontinent, there is a need to strengthen the health system. The Poly-parasitism and co-infections are major emerging problems in world of Kala azar. In recent years, the threat of HIV and Kala azar co-infection is amplifying. To date, co-infection with leishmaniasis and HIV has reached 35 countries. The risk of developing Kala azar in endemic areas is about 100 to 1000 times higher in individuals infected with HIV. HIV infection also impairs treatment response and increases the likelihood of recurrence\textsuperscript{41}. If HIV epidemic spreads to the general population where Kala azar is endemic, it may have disastrous consequences.

Kala azar may not affect the national economy or the national GDP, but it devastates the families affected. This is because the poorest of the poor who are maximally affected by the disease sink deeper into poverty following an attack of Kala azar. The strategy for elimination of Kala azar should focus intersectoral coordination aimed at improving the socio-economic status of the poorest of the poor in the community. Kala azar leads to a loss of about four lakhs DALYs every year in Nepal amounting to a loss of approximately US$ 140 million annually\textsuperscript{48}. The evidence-based policy needs to be developed to make Programme more cost-effective\textsuperscript{34}. Elimination of Kala azar in member countries of the WHO South-East Asia Region is relevant for achieving the
Millennium Development Goals (MDGs) by alleviating poverty. Policies to control Kala-azar will have to include many activities that involve public awareness and modifying behaviour. Planning of prevention and control measures for Kala-azar should take into account the possible effects of migrants and moving population. The allocated budget per person per annum for risk population by national programme for Bangladesh, India and Nepal as estimated are around 0.2, 0.4 and 0.3 US$ for Bangladesh, India and Nepal respectively\(^{34}\). Supports from international agencies and nongovernmental organizations need to be achieved to overcome the economic constrains.

Approximately 200 million people in the South East Asia Region are “at risk” of the disease, and it is now being reported from 52 districts in India, 45 districts in Bangladesh and 14 Districts in Nepal. Nepal is situated in the Southern slope of Himalaya Mountain and is surrounded by India in east, west and south and China in north. It is located between 26°22’ and 30°27’ north latitudes and 80°40’ and 88°12 east longitudes with a total surface of approximately 147181 square Km. Ecologically the country is divided from north to south into three regions: mountain, hill and terai (lowlands). The terai has a sub tropical climate and is known to harbor most tropical diseases. In Nepal, more than 6.5 million people are estimated to be at risk of VL in eastern and central Nepal bordering to Bihar, India\(^9\). In Nepal, during 2060’s and 2070’s VL ceased to be a public health problem due to blanket coverage of endemic population, by DDT to eradicate Malaria. VL cases were first reported in 1980 from Dhanusha district with the incidence rate of 1.5 per 100000 populations and case fertility rate of 5.88 percent\(^9\). Napier in 1922 reported that a large proportion of VL cases from Calcutta were Nepalese patients\(^{40}\).

In a teaching hospital 65% of the cases were found in the age group of 11 – 30 and 90% were males\(^{24}\). At national level, majority of the cases (more than 65%) were found in age group 15 years and above, with more cases reported from males, but in India nearly half of the VL cases occur in children\(^8\).

Seasonality of all the cases is not available. But whichever data is available indicates that some year it has the peak in the month of April and in the other in the month of May. Most of the cases are reported during the monsoon period (May – September) although there were cases all the year round\(^{20}\).
Distribution of Phlebotomus sand flies is highly dependent on temperature and relative humidity environment. In Southeast Asia among 115 weather stations, 62% were considered endemic with current temperature. If the temperature increased by 1° C, 12% additional station would become endemic. In Nepal, every year cases are reported from new districts, which districts is north to terai belt, one case was reported from Okhaldhunga and two cases were reported from Makawanpur districts in the year 2005. Whether the VL is increasing its area towards the hills? A qualitative survey conducted in epidemic region revealed poor knowledge about the transmission and symptoms of kala azar. Poverty, malnutrition, illiteracy and behavioral and ecological conditions are the main risk factors. Sleeping on a bed or cot and sleeping under a bed net, ever if untreated was found protective from kala azar. Ownership of cow of buffalo was protective but dampness observed in the mud floor of the house was a strong risk factor. An increase of 1 cow/sq km around the house reduce the number of indoor P argentipes in a given house by a factor of 0.92, hence it is protactive. But if one cow is owned by a person the number of indoor P argentipes in the house increase by a factor of 1.49, means it increases the risk. Similarly in mud houses risk is increased by a factor of 1.47.
Logistic regression model

Logistic regression is a type of probabilistic statistical classification model. It is also used to predict a binary response from a binary prediction, used for predicting the outcome of a categorical dependent variable based on one or more predictor variables. It is used in estimating the parameters of a qualitative response model. The probabilities describing the possible outcomes of a single trial are modeled, as a function of the predictor variables, using a logistic function. Logistic regression measures the relationship between a categorical dependent variable and one or more independent variables. Logistic regression can be binomial or multinomial. Binomial or binary logistic regression deals with situations in which the observed outcome for a dependent variable can have only two possible types (for example, "case" vs. "control"). Multinomial logistic regression deals with situations where the outcome can have three or more possible types (e.g., "disease A" vs. "disease B" vs. "disease C"). In binary logistic regression, the outcome is usually coded as "0" or "1", as this leads to the most straightforward interpretation. If a particular observed outcome for the dependent variable is the noteworthy possible outcome (referred to as a "success" or a "case") it is usually coded as "1" and the contrary outcome (referred to as a "failure" or a "noncase") as "0". Logistic regression is used to predict the odds of being a case based on the values of the independent variables (predictors). The odds are defined as the probability that a particular outcome is a case divided by the probability that it is a non case. Like other forms of regression analysis, logistic regression makes use of one or more predictor variables that may be either continuous or categorical data. Unlike ordinary linear regression, however, logistic regression is used for predicting binary outcomes of the dependent variable (treating the dependent variable as the outcome of a Bernoulli trial) rather than continuous outcomes. Given this difference, it is necessary that logistic regression take the natural logarithm of the odds of the dependent variable being a case (referred to as the logit or log-odds) to create a continuous criterion as a transformed version of the dependent variable. Thus the logit transformation is referred to as the link function in logistic regression—although the dependent variable in logistic regression is binomial, the logit is the continuous criterion upon which linear regression is conducted. The logit of success is then fit to the predictors using linear regression analysis. The predicted value of the logit is converted back into predicted odds via the inverse of the natural logarithm, namely the exponential function.
Therefore, although the observed dependent variable in logistic regression is a zero-or-one variable, the logistic regression estimates the odds, as a continuous variable, that the dependent variable is a success (a case). In some applications the odds are all that is needed. In others, a specific yes-or-no prediction is needed for whether the dependent variable is or is not a case; this categorical prediction can be based on the computed odds of a success, with predicted odds above some chosen cut-off value being translated into a prediction of a success.

**Logistic function, odds ratio, and logit**

![Logistic function graph](image)

Figure 1: the logistic function, with $\beta_0 + \beta_1 x$ on the horizontal axis and $F(x)$ on the vertical axis. An explanation of logistic regression begins with an explanation of the **logistic function**, which always takes on values between zero and one:

$$F(t) = \frac{e^t}{e^t + 1} = \frac{1}{1 + e^{-t}},$$

and viewing $t$ as a linear function of an **explanatory variable** $x$ (or of a linear combination of explanatory variables), the logistic function can be written as:

$$F(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}.$$ 

This will be interpreted as the probability of the dependent variable equalling a "success" or "case" rather than a failure or non-case. We also define the inverse of the logistic function, the **logit**:

$$g(x) = \ln \left( \frac{F(x)}{1 - F(x)} \right) = \beta_0 + \beta_1 x,$$
and equivalently:

\[
\frac{F(x)}{1 - F(x)} = e^{\beta_0 + \beta_1 x}.
\]

Multiple explanatory variables

If there are multiple explanatory variables, then the above expression \( \beta_0 + \beta_1 x \) can be revised to \( \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m \). Then when this is used in the equation relating the logged odds of a success to the values of the predictors, the linear regression will be a multiple regression with \( m \) explanatory; the parameters \( \beta_j \) for all \( j = 0, 1, 2, \ldots, m \) are all estimated.

Case-control sampling:

Suppose cases are rare. Then we might wish to sample them more frequently than their prevalence in the population. For example, suppose there is a disease that affects 1 person in 10,000 and to collect our data we need to do a complete physical. It may be too expensive to do thousands of physicals of healthy people in order to get data on only a few diseased individuals. Thus, we may evaluate more diseased individuals. This is also called unbalanced data. As a rule of thumb, sampling controls at a rate of five times the number of cases is sufficient to get enough control data.

If we form a logistic model from such data, if the model is correct, the \( \beta_j \) parameters are all correct except for \( \beta_0 \). We can correct \( \beta_0 \) if we know the true prevalence as follows:

\[
\hat{\beta}_0^* = \hat{\beta}_0 + \log \frac{\pi}{1 - \pi} - \log \frac{\tilde{\pi}}{1 - \tilde{\pi}}
\]

Where, \( \pi \) is the true prevalence and \( \tilde{\pi} \) is the prevalence in the sample.

Rational of this study:

The bordering region of Nepal and Indian state Bihar is one of the world's main endemic areas for VL, which over the last 40 years has seen a dramatic increase in the number of VL cases, due
to a complexity of factors. Several studies have convincingly shown that, data on socio-economic, behavioral and entomological risk factors for VL are comparatively insufficient, and in these areas, no study, to their knowledge, has investigated risk factors for VL. This severely hampers efforts to prevent and control the disease in the region. To fill these gaps, we will be carried out a case-control study to identify individual and household level risk factors for VL in an endemic focus in eastern terai districts of Nepal. Therefore study will be needed to bring out the important fact that kala azar is a major public health issue of South Asia region and affecting a large number of poor population. So the study will help to bring Kala azar in the national public health problem of these countries that will lead to coordinated strategies plan to eliminate the disease from this region.